Imperial College London

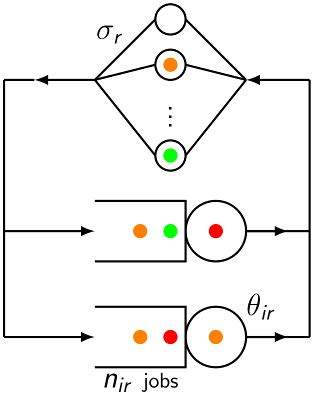
Accelerating Performance Inference over Closed Systems by Asymptotic Methods

Giuliano Casale
Imperial College London
g.casale@imperial.ac.uk

ACM SIGMETRICS 2017

Imperial College London

Closed queueing networks



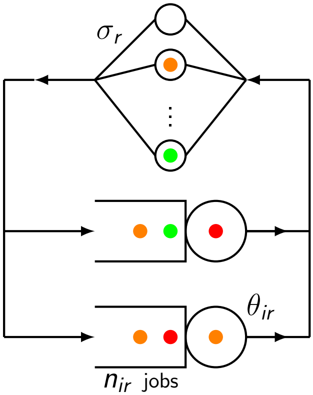
K nodes
R classes
N jobs, N_r in class r

Product-form ($\sigma_r = 0$):

$$P(\boldsymbol{n}) = \frac{1}{G} \prod_{i=1}^{K} n_i! \prod_{r=1}^{R} \frac{\theta_{ir}^{n_{ir}}}{n_{ir}!}$$

Imperial College London

Closed queueing networks



K nodes
R classes
N jobs, N_r in class r

Product-form ($\sigma_r = 0$):

$$G = \sum_{\boldsymbol{n}} \prod_{i=1}^{K} n_i! \prod_{r=1}^{R} \frac{\theta_{ir}^{n_{ir}}}{n_{ir}!}$$

See paper for case $\sigma_r > 0$.

Motivating example

- Problem: service demand inference
  - Q_{ir}: empirical mean queue-length

- Maximum-likelihood estimation (MLE)

$$\max_{\boldsymbol{\theta} \in \Omega} \sum_{i,r} Q_{ir} \log \theta_{ir} - \log G$$

where $\boldsymbol{\theta} \equiv (\theta_{ir})$ and $G \equiv G(\boldsymbol{\theta})$.

θ_{ir} service demand, N: jobs, R: classes, K: queues

Related work

Computing the normalizing constant G:

- CA: Convolution algorithm
- RECAL: Recursion by chain
- MoM: Method of moments

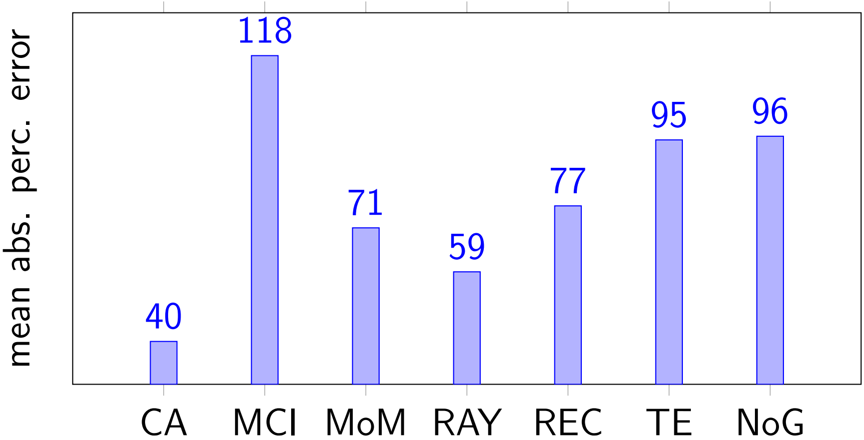
} exact methods

- RAY: Ray method
- MCI: Monte Carlo integration
- TE: Taylor expansion
- ...

Imperial College London

Motivating example

- 60s run, 3 nodes, 3 classes, 120 jobs



This prompts us to revisit G's computational theory.

Interpolation theory

- Consider $f(x)$ and data points $x = \theta_1, \ldots, \theta_K$
- Newton's interpolation polynomial

$$N(x) = \prod_{i=1}^{K}(x - \theta_i) \cdot \underbrace{[\theta_1, \ldots, \theta_K]f(x)}_{\text{divided difference}} + \cdots$$

- Explicit form for divided differences

$$[\theta_1, \ldots, \theta_K]f(x) = \sum_{k=1}^{K} \frac{f(\theta_k)}{\prod_{i \neq k}(\theta_k - \theta_i)}$$

θ_{ir} service demand, N: jobs, R: classes, K: queues

Single-class models

- Explicit form for single-class models ($R = 1$)

$$G = \sum_{\boldsymbol{n}} \prod_{k=1}^{K} \theta_k^{n_k} = \sum_{k=1}^{K} \frac{\theta_k^{N+K-1}}{\prod_{i \neq k}(\theta_k - \theta_i)}$$

- Matching the definitions

$$G = [\theta_1, \ldots, \theta_K] x^{N+K-1}$$

θ_{ir} service demand, N: jobs, R: classes, K: queues

A new integral form

- Hermite-Genocchi formula for divided differences

$$G = \frac{(N+K-1)!}{N!} \int_{\Delta_K} (\theta_1 u_1 + \ldots + \theta_K u_K)^N d\boldsymbol{u}$$

  on the simplex $\Delta_K = \{|\boldsymbol{u}| = 1, \boldsymbol{u} \geq \mathbf{0}\}$.

- Laplace transform on volume of Δ_K yields the McKenna-Mitra integral form for G over $\mathbb{R}_+^K$.

Imperial College London

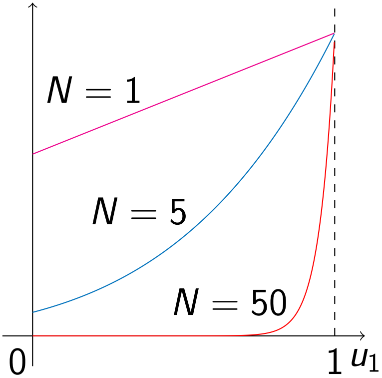
Example

$K = 2$ queues

$u_1 + u_2 = 1$

$\theta_1 > \theta_2$

$$G \propto \int_{\Delta_K} (\theta_1 u_1 + \theta_2 u_2)^N d\boldsymbol{u}$$



Imperial College London

Multiclass analysis

- We generalize the result to multiclass models

$$G = \frac{(N+K-1)!}{N_1!\cdots N_R!}\int_{\Delta_K}\prod_{r=1}^{R}(\theta_{1r}u_1+\ldots+\theta_{Kr}u_K)^{N_r}d\boldsymbol{u}$$

- Proof uses multinomial theorem and

$$P(\boldsymbol{n}) \propto n_1!\cdots n_K! \propto \int_{\Delta_K}\prod_{k=1}^{K}u_k^{n_k}d\boldsymbol{u}$$

θ_{ir} service demand, N: jobs, R: classes, K: queues

Imperial College London

Computational methods

From the integral form, we derive:

- Explicit solutions
- Asymptotic expansion
- Cubature rule

Computational methods

From the integral form, we derive:

- Explicit solutions
- Asymptotic expansion
- Cubature rule

Explicit solutions

- Integral form for single class models ($R = 1$)

$$G = \frac{(N+K-1)!}{N!} \int_{\Delta_K} (\theta_1 u_1 + \ldots + \theta_K u_K)^N d\boldsymbol{u}$$

- Integral form for multiclass models ($R > 1$)

$$G = \frac{(N+K-1)!}{N_1! \cdots N_R!} \int_{\Delta_K} \prod_{r=1}^{R} (\theta_{1r} u_1 + \ldots + \theta_{Kr} u_K)^{N_r} d\boldsymbol{u}$$

θ_{ir} service demand, N: jobs, R: classes, K: queues

From products to sum

- Can we relate the two cases?

$$\prod_{r=1}^{R} x_r^{N_r} = \frac{1}{N!} \frac{\partial^{N_1} \dots \partial^{N_R}}{\partial t_1^{N_1} \dots \partial t_R^{N_R}} \left(\sum_r t_r x_r \right)^N$$

where $N = \sum_r N_r$.

- Multiclass integrand reduced to single class!

$O(N^R)$ Explicit form

- Replace derivatives with finite differences.
- Applying to the integrands

$$G = \sum_{\boldsymbol{t}:0 \le t_r \le N_r} \underbrace{\prod_{r=1}^{R} \frac{(-1)^{N_r - t_r}}{N_r!} \binom{N_r}{t_r}}_{\text{finite difference}} \underbrace{\sum_{k=1}^{K} \frac{\theta_k^{N+M-1}}{\prod_{i \neq k}(\theta_i - \theta_k)}}_{\text{single-class solution}}$$

where $\theta_k = \sum_r t_r \theta_{kr}$.

- Same time complexity as CA, but $O(1)$ space

$O(N^{K+1})$ Explicit form

- Applying the idea to products in $P(\boldsymbol{n})$

$$G = \sum_{\boldsymbol{h}\geq\mathbf{0}:h\leq N} \frac{(-1)^{N-h}}{N_1!\cdots N_R!}\binom{N+K-1}{N-h}\prod_{r=1}^{R}\left(\sum_{k=1}^{K} h_k\theta_{kr}\right)^{N_r}$$

- Same time as RECAL, but $O(1)$ space

θ_{ir} service demand, N: jobs, R: classes, K: queues

Computational methods

From the integral form, we derive:

- Explicit solutions
- Asymptotic expansion
- Cubature rule

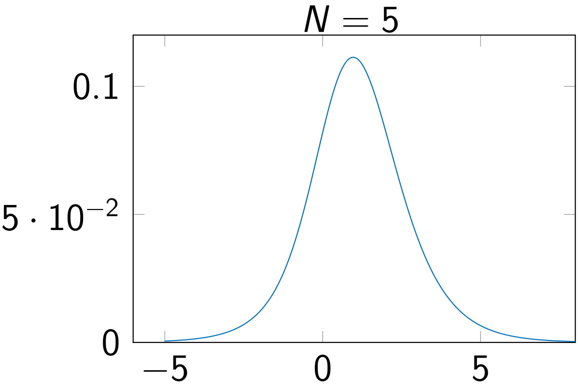
Imperial College London

Logistic transformation

$$u_1 \to \frac{e^x}{1+e^x}$$

$$u_2 \to \frac{1}{1+e^x}$$

$$\left|\frac{\partial u}{\partial x}\right| = \frac{e^x}{(1+e^x)^2}$$



- Approximate mapping to a Gaussian over $\mathbb{R}^{K-1}$
- Close to Gaussian also for small N

Imperial College London

Properties

- What are the parameters of the Gaussian?
- Mean $\hat{\boldsymbol{u}}$ obtained iteratively

$$\hat{u}_i = \sum_{r=1}^{R} \frac{N_r}{N \sum_k \theta_{kr} \hat{u}_k}$$

- Asympotically identical to mean-value analysis
- Explicit form found for the covariance matrix $\boldsymbol{A}$

θ_{ir} service demand, N: jobs, R: classes, K: queues

Heavy-load asymptotics

- We force ϵN self-looping jobs at each node
- For small ϵ, we get in heavy load ($N \to \infty$)

$$G \sim \frac{(N+K-1)!}{N_1! \cdots N_R!} \sqrt{\frac{(2\pi)^{K-1}}{\det(\boldsymbol{A})}} \prod_{r=1}^{R} \left(\sum_{k=1}^{K} \theta_{kr} \hat{u}_k \right)^{N_r} \prod_{i=1}^{K} \hat{u}_i$$

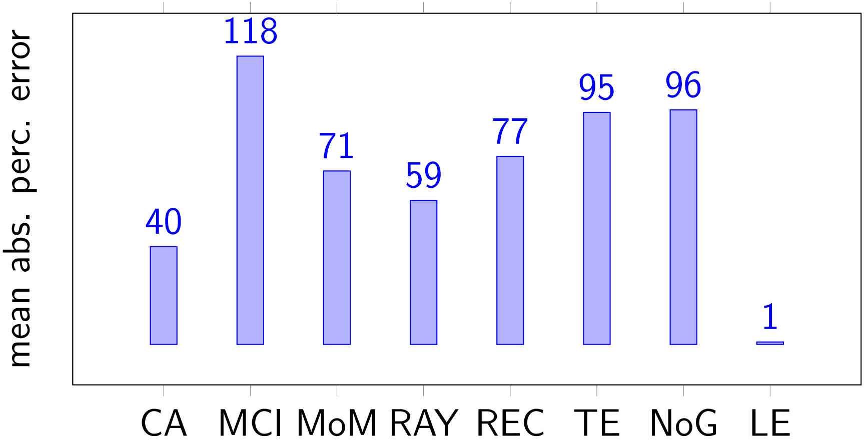
- We refer to this result as logistic expansion (LE).

θ_{ir} service demand, N: jobs, R: classes, K: queues

Imperial College London

Example: MLE revisited

- 60s run, 3 nodes, 3 classes, 120 jobs

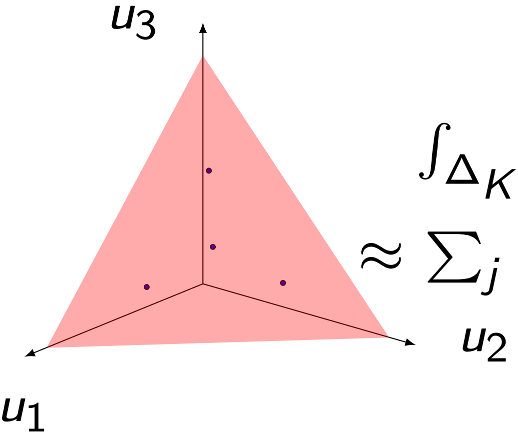


Computational methods

From the integral form, we derive:

- Explicit solutions
- Importance sampling
- Asymptotic expansion
- Cubature rule

Cubature rules



$$\int_{\Delta_K} f(\boldsymbol{u})d\boldsymbol{u} \approx \sum_j w_j f(\boldsymbol{u}^{(j)})$$

Grundmann-Möller cubature

- $O(S^K)$ samples symmetric to simplex barycenter

$$\boldsymbol{u}^{(j)} = \left(\frac{(2b_1^{(j)} + 1)}{(2S + K - 2s)}, \ldots, \frac{(2b_K^{(j)} + 1)}{(2S + K - 2s)} \right)$$

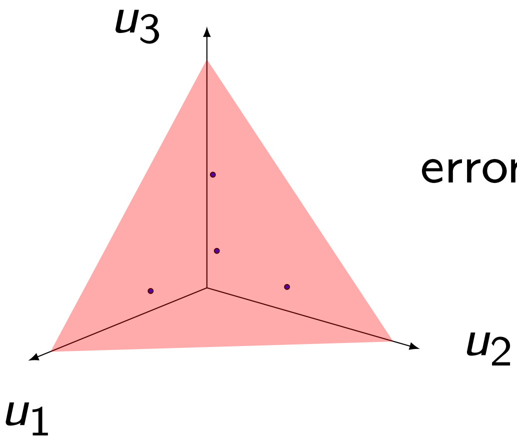
$$\forall s = 0, \ldots, S \text{ and } \boldsymbol{b}^{(j)} \geq \mathbf{0} \text{ s.t. } |\boldsymbol{b}^{(j)}| = S - s.$$

- For G, the rule is exact when $S = \lceil (N-1)/2 \rceil$
- Lower degree rules give approximations of G

Grundmann-Möller cubature

degree $S = 1$

4 points

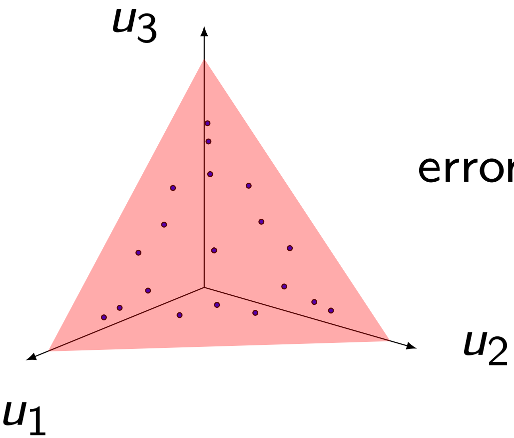


error=92.9%

Grundmann-Möller cubature

degree $S = 2$

18 points



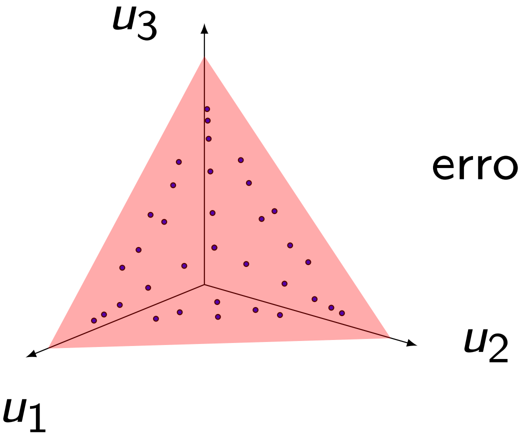
error=61.0%

Imperial College London

Grundmann-Möller cubature

degree $S = 4$

35 points

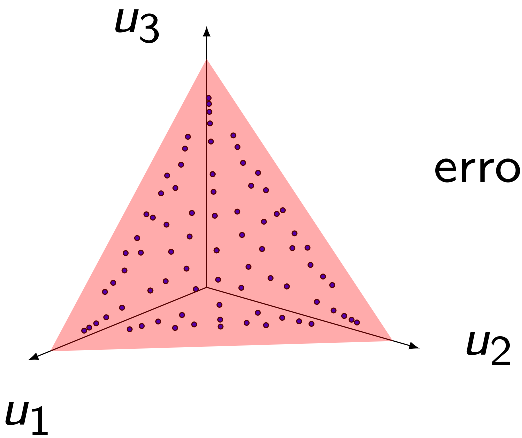


error=6.0%

Grundmann-Möller cubature

degree $S = 6$

84 points

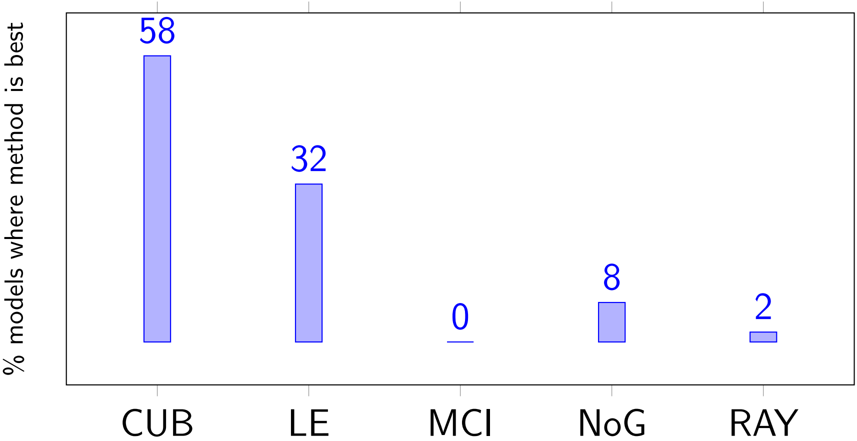


error=0.0%

Imperial College London

Validation: Light-load

- $K, R \in [2, 32]$, $N/R = 2$, 1125 MLE problems

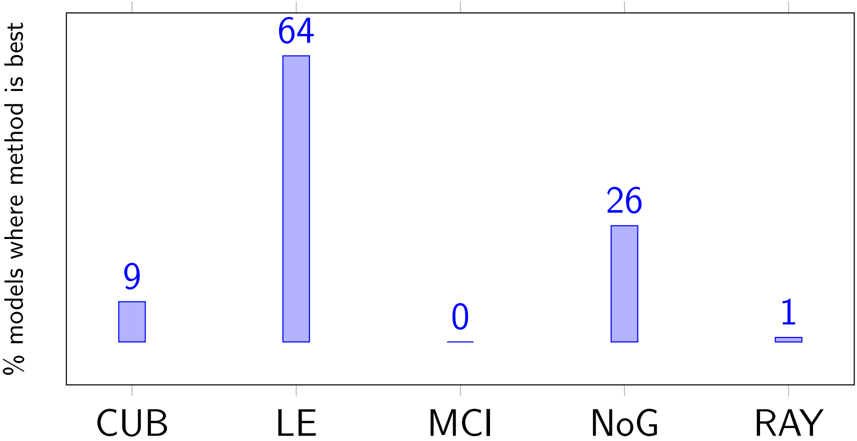


Cubature rule best in light load.

Imperial College London

Validation: Heavy-load

- $K, R \in [2, 32]$, $N/R = 40$, 1125 MLE problems



Logistic expansion best in heavy load.

Conclusion

Main findings:

- Novel integral form for G on the unit simplex
- Explicit solutions
- Asymptotic expansions and cubature rule

Additional results in the paper:

- Importance sampling method
- Models with delay nodes ($\sigma_r > 0$)
- Validation on random models